CAUSAL ASSUMPTIONS:

SOME RESPONSES TO NANCY CARTWRIGHT

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Computer Science

University of Saskatchewan

Saskatoon

By

SONJE KRISTTORN

Keywords: causal inference, causal markov condition, faithfulness assumption, causal graphs

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
176 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
S7N 5C9

ABSTRACT

The theories of causality put forward by Pearl and the Spirtes-Glymour-Scheines group have entered the mainstream of statistical thinking. These theories show that under ideal conditions, causal relationships can be inferred from purely statistical observational data. Nancy Cartwright advances certain arguments against these causal inference algorithms: the well-known "factory example" argument against the Causal Markov condition and an argument against faithfulness. We point to the dependence of the first argument on undefined categories external to the technical apparatus of causal inference algorithms. We acknowledge the possible practical implication of her second argument, yet we maintain, with respect to both arguments, that this variety of causal inference, if not universal, is nonetheless eminently useful. Cartwright argues against assumptions that are essential not only to causal inference algorithms but to causal inference generally, even if, as she contends, they are not without exception and that the same is true of other, likewise essential, assumptions. We indicate that causal inference is an iterative process and that causal inference algorithms assist, rather than replace, that process as performed by human beings.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Eric Neufeld, for help and encouragement, always kindly and generously given. I am also grateful for the companionship and support of my friends, Qingjuan Guan and Maryruth Pradeepa Joseph Lawrence, inside and outside the lab. My thanks, too, to the members of my Advisory Committee, who read the thesis and offered helpful comments and advice.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Certain algorithms, it is claimed, can derive causal structure form observational data. Perhaps the best known of these are the *IC* (*Inductive Causation*) algorithm put forward by Judea Pearl and Thomas Verma [22] [23] and the *PC* (probably named for *Peter* and *Clark*) algorithm proposed by Peter Spirtes, Clark Glymour, and Richard Scheines (*SGS* hereafter) [27]. The IC algorithm assumes perfect information and so abstracts from problems of data and error. As a result, it has an uncluttered clarity that is well suited to our purposes and we use it as our paradigmatic causal inference algorithm in what follows.

There are some who find the claims made for these algorithms "too good to be true". A prominent critic along these lines is philosopher Nancy Cartwright. In a series of books and articles she has argued against causal inference algorithms. Our discussion centres on two arguments in *The Dappled World* [3]. Broadly speaking, she argues that causal inference algorithms are valid only in restricted circumstances and that, in fact, the world is far more complex and untidy than they assume. More precisely, she challenges two assumptions fundamental to causal inference algorithms. The first, in informal terms, is the assumption that "dependency implies causality". This is not the naïve media fallacy that correlation (or dependency) indicates a cause and effect relationship but a deeper notion that the presence of dependencies manifests an underlying, explanatory causal structure. We call this the *Fundamental Assumption* and we argue that it is the real target of her *factory example* argument. The second is the assumption of *Faithfulness*, that independencies and dependencies in data faithfully reflect causal structure.

In response, again speaking broadly, we note that many of the complicating features of the world she alleges, features, she claims, that render the Fundamental Assumption invalid in many if not most instances, lack formal or precise definition, making it difficult to assess or respond to her arguments. Further, we observe that reasoning in a messy, complicated world relies on general principles that are not universally valid but essential nonetheless. Closely tied to this is another observation, that causal inference, like other forms of reasoning, is an iterative process. We maintain that if the two assumptions under discussion are not without exception, they fall into the category of fallible but necessary assumptions, and we point out that causal inference algorithms are not meant to replace human reasoning but to assist it, to serve as a tool in the iterative process of causal inference. However useful the tool – and we regard causal inference algorithms as exceptionally useful – human intervention remains critical and the fact that reasoning about causality, like other forms of reasoning, depends on general principles or assumptions that lack universal validity is one reason why this is so. We illustrate this with respect to the Faithfulness Assumption.

In the next chapter (chapter 2), we review relevant notation, definitions, assumptions, and a simple version of the IC algorithm. With this foundation in place, chapter 3 presents Cartwright's arguments. Chapter 4 contains our responses to those arguments. In the last chapter, we summarize our conclusions.

CHAPTER 2
NOTATION AND TERMINOLOGY, DEFINITIONS, ASSUMPTIONS, AND ALGORITHMS

Before we consider Cartwright's arguments, we need to give an account of what she is criticizing. What follows is by no means complete, but it is sufficient for our purposes and omits unnecessary, and perhaps distracting and confusing, features.

DAGs and Probability Distributions

Call $S = \{x_1, \ldots, x_n\}$, a set of individuals under discussion, a *sample space.* Then a variable A is any partition of S. For the present discussion, A divides S into two parts, which we typically identify as a and $\sim a$, the *outcomes* of A. Here, A is a binary variable.[1] Then, we can say that

$$p(a) = \frac{\| a \|}{\| S \|} \tag{1.1}$$

The joint probability $p(ab)$ denotes the proportion of individuals in both a and b, and can be written

$$p(ab) = \frac{\| a \wedge b \|}{\| S \|} \tag{1.2}$$

Where $a \wedge b$ is shorthand for the set of elements of S that are members of both a and b. Then, when $p(b) \neq 0$, the conditional probability $p(a|b)$ is just

$$p(a \mid b) = \frac{p(ab)}{p(b)} \tag{1.3}$$

[1] It is easy to generalize the discussion to *n*-ary variables, including continuous variables, but because that would require discussion and notational machinery superfluous to our purposes, we do not do so.

Variables A, B are conditionally independent given a set of variables C if

$$p(A,B|C) = p(A|C) \cdot p(B|C) \tag{1.4}$$

for all outcomes of A, B and C. If C is empty, then A, B are unconditionally independent.

A *graph* is a pair $G=(V,E)$ where V is a set of vertices (or nodes) and E is a set of edges or arcs between pairs of vertices (v_n, v_m). In a *directed graph*, the edges have direction, so that the direction of an arc (v_n, v_m) is from v_n to v_m. A cycle is a path, following the directionality of the arcs from one vertex to another, that begins and ends at the same vertex. Thus, a cycle C_n consists of n vertices $v_1, v_2, ..., v_n$ connected by edges $(v_1, v_2), (v_2, v_3), ..., (v_{n-1}, v_n), (v_n, v_1)$. A *Directed Acyclic Graph* (*DAG*) is a directed graph that has no cycles.[2]

Our concern is with DAGS as they are used to represent probability distributions (although they have other applications, too). A DAG that represents a probability distribution is a pair (D,P), where $D=(V,E)$ is a directed graph, vertices represent variables, and P is a probability distribution over V. The structure of the graph represents independence – and, conversely, dependence – relationships in the data as follows: For any variable A, variables directly connected to A by incoming arcs are denoted *parents*(A), and described as the parents of A. A itself is a *child* of its parents and parent of its children, themselves *descendents* of the parents of A. A is conditionally independent of all other variables, its descendants excepted, given its parents. Associated with A is a local distribution $f(A, \mathit{parents}(A))$, which gives a distribution for A for any set of values that *parents*(A) take on. f can be a discrete or continuous probability distribution or a

[2] In causal graphs, acyclicity corresponds to the intuition that an event can't cause itself, though we are running somewhat ahead of ourselves in making this observation.

structural equation.[3] The structure of the graph thus encodes the information that *P* factors into the product of the conditional distributions stored at the nodes. For discrete variables, then,

$$P(v_0,\ldots,v_n) = \Pi_i\, P(v_i|parents(v_i)) \quad (1.5)$$

Factorization can make probabilistic reasoning computationally feasible in both time and space, where the graph is sparse. This way of representing independence relationships in the data is summarized in what is termed the *Markov Condition* (*MC*): A vertex is conditionally independent of its non-descendants, given its parents.[4]

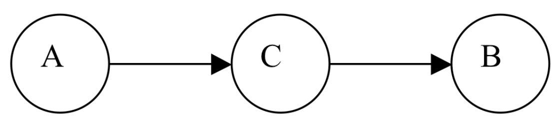

Figure 2.1: B is conditionally independent of A, given C

For example, the simple graph in Figure 2.1 encodes the information that B is independent of its non-descendent, *A*, given its parent, *C*. *A* and *B* are dependent, but not dependent in all conditioning contexts. On the other hand, a node and its parent(s), are always dependent, that is, they are dependent in all contexts, for all conditioning sets. So the graph shows that *A*, *C* and *C*, *B* are always dependent or dependent in all contexts.

[3] Here, we focus on discrete distributions, though the algorithms and definitions discussed employ an abstract notion of independence.
[4] The Markov Condition implies many other conditional independencies. These can be detected from *D* using a graph-theoretic criterion called *d-separation* [23]. But we need not discuss this.

Note that the graph in Figure 2.2 expresses the same set of independencies (and, therefore, dependencies).

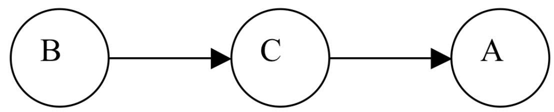

Figure 2.2: A is conditionally independent of B, given C

In Figure 2.3, *A* and *B* have a common parent, *C.* The graph shows that *A*, *C* and *B*, *C* are always dependent. It also indicates that *A* is conditionally independent of its non-descendent *B*, given its parent, *C*, and equivalently, that *B* is conditionally independent of *A*, given *C*. Otherwise, *A* and *B* are dependent. This is the same set of independence relations portrayed both in Figure 2.1 and its mirror image, where the directions of the arcs are reversed, in Figure 2.2. Often a given set of independencies can be represented by different graphical structures. Graphs that represent the same independence relations are said to be *Markov equivalent*.

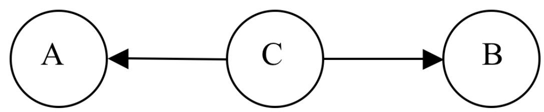

Figure 2.3: A is conditionally independent of B, given C

While the three graphs shown above are Markov equivalent, the *head-to-head* structure in Figure 2.4 shows that *A* and *C* are always dependent, as are *B* and *C*. *A,* which has no parents*,* is unconditionally independent of its non-descendent *B* and *B* is likewise unconditionally independent of *A*.[5] On the other hand, *A* and *B* are conditionally dependent, given *C*. If this last seems counter-intuitive, consider the dependencies between *A*, *C* and *B*, *C*. Knowing the value of *A* tells us nothing about the value of *B*. However, if we know the value of *C*, the value of *A* will tell us something about the value of *B*, since *C* is dependent on both *A* and *B*. Suppose, for example, that *Cold* (*A*) and *Sneeze* (*C*) are always dependent and that *Hayfever* (*B*) and *Sneeze* (*C*) are always dependent, while *Cold* (*A*) and *Hayfever* (*B*) are independent. The fact that someone has or does not have a *Cold* (*A*) has no bearing on the probability that the same person has *Hayfever* (*B*). However, if we know that someone is *Sneezing* (*C*), then knowing that he or she does not have a *Cold* (*A*) *does* tell us something about the probability that the

[5] It is not necessary for the unconditional independence of A and B that A and B be ‘parentless’ but to avoid unnecessary complications, and because in the example A and B are orphans, we stated the matter as we did.

person has *Hayfever* (*B*). Once we know the value for *Sneezing* (*C*), *Cold* (*A*) and *Hayfever* (*B*) are dependent.

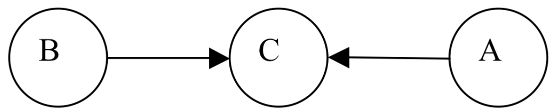

Figure 2.4: A and B are conditionally dependent, given C

The MC implies that all independencies represented in the graph exist in the data the graph represents. But it does not imply that all independencies in the data can be derived from the graph. To put it another way, the MC does not say that the distribution contains only the independencies represented in the graph or, to put it yet another way, that all independencies in the data are represented in the graph. That information is given in a version of the *Faithfulness Assumption* (*FF*). We refer to this version as *Faithfulness1*(*FF1*) to distinguish it from another, and much more common, version discussed below, usually referred to as the *Faithfulness Assumption* without qualification. The former, FF1, unlike FF, is not an assumption at all but a further description of the relationship between a probability distribution and a DAG. It says that all independencies in the data are represented in the graph. As such, it supplies the converse of the *if* contained in the Markov Condition. The upshot is an *if and only if* between the independencies represented in a graph and independencies in a distribution. According to the MC, *if* an independence relationship is represented in the graph, *then* it is present in the data. FF1 says that *if* an independence relationship is present in the data, *then* it is represented in the graph.

Causal Graphs

In the previous section, a DAG was presented as a compact encoding of a joint probability distribution that exploits independence knowledge to gain efficiency in space and time. In this section, we consider DAGs as representative of causal relationships between variables.[6] The algorithms and definitions in this section all assume the availability of *perfect* (zero sampling error) measurements. We also assume that our domain has no cycles or feedback loops.[7] Again, for the most part, the only association mentioned is probabilistic dependence, though the definitions and algorithms can easily accommodate correlations and covariances between continuous variables.

A causal interpretation of a DAG that represents a probability distribution depends on two assumptions and some definitions. Both assumptions concern the relationship between probability distributions and causal relationships. The first assumption is that dependency implies causation: if *A* and *B* are dependent in all contexts, then *A* causes *B* (directly, though there may be unmeasured mediating causes), *B* causes *A* (again directly, with the same qualification), or *A* and *B* have a hidden common cause which would, if measured and conditioned on, render them conditionally independent. This also means (by *modus tollens*) that the absence of these three causal relationships between *A* and *B* implies independence in some context (including the empty context).

[6] We shall employ a common verbal shortcut and speak of causal relationships between variables rather than, more precisely, of causal relationships between what the variables represent.
[7] In fact, it is possible to take feedback into account, but this simplified presentation omits that possibility.

We'll call this the *Fundamental Assumption* (*FA*). Informally, it amounts to the claim that dependency implies causality and that absence of causality implies independence.

The second, *faithfulness* [28] (or *stability* [23]) assumption, is that all independencies in the distribution are entailed by causal structure. To paraphrase Kyburg [15], no independencies result from "numerical coincidences". (This is what, above, we called *FF*.) There is a close relationship between the Faithfulness assumption and the Fundamental Assumption. Essential to FF is an idea that is the converse of FA: if *A* causes *B*, *B* causes *A*, or *A* and *B* have a hidden common cause, *A* and *B* will not be independent in any context, that is, they will be dependent in all contexts. Equivalently, if *A* and *B* are independent in some context, including the empty context, then *A* does not cause *B* (directly, as above), *B* does not cause *A* (directly, likewise as above), and *A* and *B* do not have a hidden common cause. More succinctly, and less precisely, independence implies no causation and causation implies dependence.

Faithfulness says that there are no independencies that do not reflect causal structure, no independencies between variables that have a (direct) causal relationship. Suppose, for example, that playing a musical *Instrument* (*I*) reduces *Stress* (*S*). But suppose also that the same activity leads to *Conflict* (*C*) with neighbours or housemates and that *Conflict* increases *Stress*. Figure 2.5 depicts this causal structure. What would happen if the amount of stress reduced by the direct causal pathway $I \rightarrow S$ were exactly equal to the stress created on the causal pathway $I \rightarrow C \rightarrow S$? In that case, $p(S|I)$ would equal $p(S)$; that is, *S* and *I* would be unconditionally independent in the distribution in spite of the direct causal relationship between them. The Faithfulness Assumption assumes that this kind of thing does not happen; it assumes that dependencies along different causal paths do not exactly balance or cancel each other out so as to create an apparent independence that fails to reflect the true causal relationship between variables. The Faithfulness Assumption might be considered an application of Ockham's razor to causal inference. Failure to make this assumption, on this understanding, introduces unnecessary complexity.

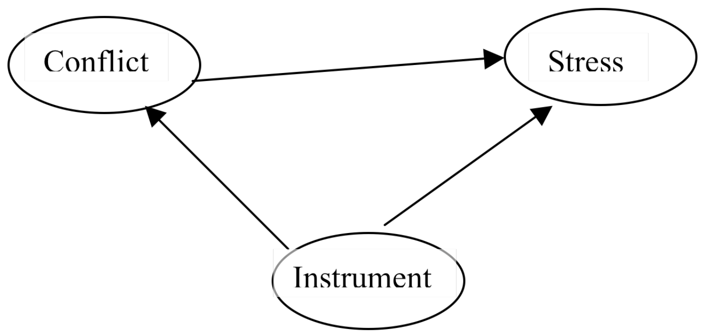

Figure 2.5: Causal relationships between playing an *Instrument*, *Conflict*, and *Stress*

The import of the Fundamental Assumption and the Faithfulness Assumption taken together is that if variables are dependent in all contexts, there is a direct causal relationship between them or a hidden common cause (FA) and if there is a direct causal relationship between variables or a hidden common cause, they will be dependent in all contexts (contra positive of FF and converse of FA). Likewise, if there is no direct causal relationship, the variables will be independent in some context (contra positive of FA) and if two variables are independent in some context, then there is no direct causal relationship between them (FF). Hence, there is dependency in all contexts if and only if there is a direct causal relationship between variables or a hidden common cause and there is independency in some context if and only if there is no direct causal relationship or hidden common cause[8].

Judea Pearl's definition of *potential cause* (*Definition 1*) [22], builds on these assumptions:

A is a potential cause of *C* if there is a variable *B* and a context *S* such that

1. *A*, *B* are independent given *S*,
2. there is no context *T* such that A and C are conditionally independent given *T*, and
3. *B* and *C* are dependent.

Condition 1 says that there is a variable or set of variables, *S*, including the empty set, that, when conditioned on, renders *A* and *B* conditionally independent. If *S* is the empty set, *A* and *B* are unconditionally independent but conditional independence is enough to satisfy condition 1. Condition 2 says that *A* and *C* are always dependent, that is, that they are dependent in all contexts. Condition 3 requires that *B* and *C* are dependent. This is consistent with their being independent in some context or, like *A* and *C*, dependent in all contexts.

A simple example illustrates the intuition behind this definition. The example satisfies the above definition, with some additional information. A and B are unconditionally independent (*S* is the empty set) or, if they are conditionally independent, *S* does not include *C*. Moreover, B and C are dependent in all contexts. In the example, then, the definition applies in two ways: First, *A* is potential cause of *C*, where the dependency between *A* and *C* satisfies condition 2 and the dependency between *B* and *C*

[8] This is sometimes given a more relaxed form: Dependence means causality and independence means no causality.

satisfies condition 3. Secondly, *B* is a potential cause of *C*, where the dependency between *B* and *C* satisfies condition 2 and the dependency between *A* and *C* satisfies condition 3.

Figure 2.6 shows all possible orientations of the arcs between *A*, *B*, and *C*. On the basis of the representation discussed in section 1 above, summarized in the Markov condition, only the fourth graph is consistent with the dependencies and independencies specified by the definition (with the additional restrictions). In the first graph, *B* is conditionally independent of its non-descendant, *A*, given its parent, *C*, and otherwise dependent, contrary to the above information, according to which *S* (condition 1) is the empty set (so that *A* and *B* are unconditionally independent) or *S* does not include *C*. The same reasoning applies to its mirror image in the second configuration. In the third graph, *A* and *B* are again dependent and conditionally independent, given *C*, and, again, this is inconsistent with the information about the relationships between the variables. Only the orientation of the arcs in the fourth image represents the dependencies and independencies specified in the (more restrictive) definition: *A* and *B* are unconditionally independent while *A*, *C* and *B*, *C* are dependent in all contexts.

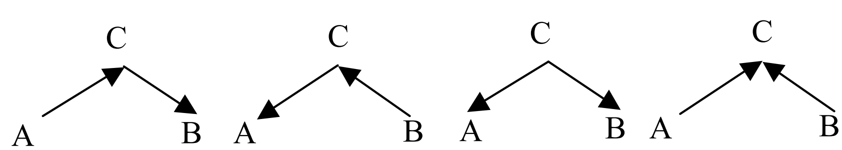

Figure 2.6: Possible orientations of the arcs

Thus far, all we have done is mapped the dependencies and independencies given in the (restricted) definition. If we invoke the relationships between dependency and causality specified in our two assumptions, we can give the four graphs a causal interpretation. By those assumptions, dependency in all contexts exists if and only if there is a direct causal relationship between variables or a hidden common cause and independency in some context (including the empty set) if and only if there is no direct causal relationship.

So the first two graphs, which are ruled out because they are inconsistent with the dependencies given in the (restricted) definition, represent, on these assumptions, a causal chain from *A* to *C* and from *C* to *B* or, in the other direction, from *B* to *C* and *C* to *A*. According to the third, *C* is a common cause of *A* and *B*. The fourth, the only one consistent with the definition, indicates *A* and *B* as causes of *C*.

In fact, the last statement needs to be qualified. The fourth graph shows that *A* and *C* and *B* and *C* are always dependent and that *A* and *B* are independent. According to the Fundamental Assumption, this means that *A* causes *C* *or* that *A* and *C* have an unmeasured common cause. The same is true of *B* and *C*. For this reason, the definition is a definition of *potential* cause.

Where certain conditions are met, it is possible to exclude the possibility of a hidden common cause. These conditions are given in Pearl's definition of *Genuine Cause* [22] (*Defintion 2*). A variable *C* is a genuine cause of *D* if

1. *A* is a potential cause of *C*.
2. *C* and *D* are dependent in all contexts.
3. *A* and *D* are dependent.
4. *A* and *D* are conditionally independent given *C*.

Figure 2.7 depicts these relations, using a bold arrow to indicate genuine causality:

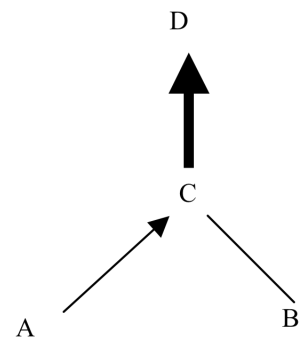

Figure 2.7: C is a genuine cause of D

This definition depends on the idea that the dependence between *C* and *D* cannot be explained by a hidden common cause (*U*) without coming into conflict with the independence information in the data. This is easiest to see if we suppose that *U* is measured and represent the dependencies/independencies involved in a graph. According to our assumptions about causality and dependence, if *U* is a common cause of *D* and *C* then *U*, *C* and *U*, *D* are dependent in all contexts and *C* and *D* are conditionally independent, given *U*. The corresponding graph would look like the one in Figure 2.8, where U is represented by a question mark. It is apparent from the graph, using the MC to see the independencies involved, that *A* and *D* are unconditionally independent and conditionally dependent given *C* contrary to the conditions set forth in the definition.

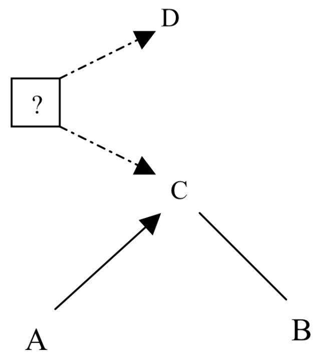

Figure 2.8: A hidden common cause is inconsistent with the definition

A non-causal version of this definition makes no assumptions about the relationship between causality and dependency. Rather, it shows that the possibility that the dependence between two variables derives from their common dependence on a third unmeasured variable can be eliminated where the relationships given in the definition obtain. (Condition 1 refers to potential cause but need not. As we saw in our earlier discussion of the definition of Potential Cause, a non-causal version of this definition can be given in terms of dependence and independence relationships.) The graph in Figure 2.8 depicts this situation – *U* and *C* always dependent, *U* and *D* always dependent, *C* and *D* conditionally independent, given *U* – and, as above, shows its incompatibility with the dependencies and independencies given in the definition. This is a useful result.

It is also possible in certain cases to detect the presence of a hidden common cause or to detect the presence of a hidden variable. Taking the causal version first, where there are two variables, each of which is a potential cause of the other, the explanation is a hidden common cause. Thus, two variables, *B* and *C*, have a *hidden common cause* (*Definition 3*) if

1. *C* and *B* are always dependent.
2. *A* and *C* are always dependent.
3. *D* and *B* are always dependent.
4. *D* and *C* are unconditionally independent.
5. *D* and *C* are dependent given *B*.
6. *A* and *B* are unconditionally independent.
7. *A* and *B* are dependent given *C*.

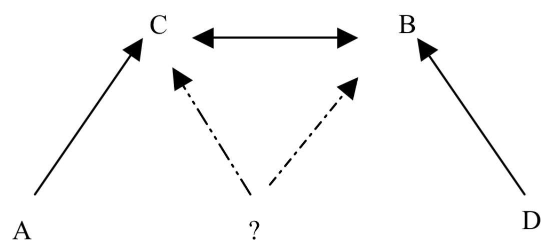

Figure 2.9: *Y* and *C* have a hidden common cause

These conditions are represented in a DAG as shown in Figure 2.9. If we interpret the graph causally, in accordance with our assumptions, the double-headed arc between C and B means either that C causes B *and* B causes C or that C and B have a hidden common cause. Since it cannot be that *C* causes *B* and *B* causes *C* (at the same time), they must have a hidden common cause.

Another way to understand this is to consider that *A*, *B*, and *C* satisfy the definition of potential cause given above (Definition 1) so that *B* is a potential cause of *C*. (*A* is also a potential cause of C but that is beside the point here.) Similarly, *C*, *B*, and *D* satisfy the definition of potential cause. *B* is therefore a potential cause of *C* (and *D* of *B* but, again, this is unimportant). This makes *C* and *B* each a potential cause of the other. Since they cannot be genuine causes of each other, and since the definition of potential cause includes the possibility of a hidden common cause, it must be that *C* and *B* have a hidden common cause.

The non-causal version of the definition determines that the dependency between two variables derives from their common dependence on a third, unmeasured variable. The graph that corresponds to Definition 3 has head-to-head arcs at *C* and *B*. (Alternatively, we can say that, according to Definition 3, *A*, *C*, and *B* satisfy Definition 1 as do *C*, *B*, and *D*.) The difficulty lies with the resulting double-headed arc between *C* and *B* which indicates that *B* is conditionally independent of *A*, given *C* (*A* -> *C* -> *B*), contrary to conditions 5 and 6. Likewise, *C* is conditionally independent of *D*, given *B* (*C* <- *B* <- *D*) contrary to conditions 4 and 5. The solution is to replace the double-headed arc with a latent (that is, a variable that has not been measured or represented) variable as parent of *C* and *B*. This representation both eliminates and explains the contradictions involved in the representation it replaces.

Pearl's causal graph construction algorithm, the *Inductive Causation* (*IC*) algorithm, builds on these definitions. Here is a simple version of the algorithm, based on the generalized algorithm in [23]:

1. For each pair of vertices *A*, *B* in a set of vertices, *V*, search for a subset *S* of *V* (including the empty set) such that *A* is conditionally independent of *B* given *S*. If no such *S* exists, let there be an undirected edge between *A* and *B*.

2. For each collinear triple of vertices *A*—*C*—*B*, where *A* and *B* are nonadjacent, test whether the *S* that renders *A*, *B* independent includes C. If not, add head-to-head arrows at C. Repeat this step until no more arrows can be added.

3. For all remaining undirected arcs, add arrows, subject to two constraints: Do not create any new structures like those in Step 2 and do not create any directed cycles. Given the results of the previous steps and subject to these constraints, generate all possible graphs. Then take the intersection of those graphs so that any arc that is consistently oriented in all the graphs is so oriented in the output graph; otherwise, the arc is unoriented.

Initially, every node is connected to every other node by an undirected link. In step one, for each pair of nodes A, *B,* the algorithm searches for a (possibly empty) set of nodes *S* such that *A* and *B* are independent given *S*. If such a set *S* exists, the link between *A* and *B* is removed. This step determines all direct links, which show that the variables concerned are always dependent, that is, they are not independent in any context.

In step two, the algorithm looks at each collinear triple, *A*—*C*—*B*. The direct links indicate that *A*, *C* are always dependent and that *C*, *B* are always dependent and the lack of a direct link between *A*, *B* shows that they are conditionally or unconditionally independent, given some *S* (empty if they are unconditionally independent). Step two checks that *S* does not include *C* and then orients the arcs in the only way that is consistent with these dependencies and independencies. (See the discussion of Definition 1 above.)

Step three is logical rather than probabilistic. Any structures like those determined in step two will have been found in step two. In addition, it is assumed that the graph does not contain directed cycles. Finally, only orientations that would appear in any graph generated by the algorithm appear in the final output.

The algorithm as presented above refers to dependencies and independencies in the data but makes no reference to causality. However, invoking our assumptions about dependency and causality, we can give a causal interpretation both to the algorithm itself and to the resulting graphs. In step one, two variables are joined by an arc if they are always dependent. On our two assumptions, such a dependency exists if and only if there is a causal relationship between the two variables. Either one is the cause of the other or they have a hidden common cause. Step two determines that A and B are potential causes of C and orients the arcs accordingly, as described earlier. The further orientation of the arcs that takes place in step three provides further indications of causal direction.

The output of the algorithm is a graph containing undirected arcs, directed arcs, and arcs directed in both directions. Directed arcs indicate not only dependency relationships between variables, as described in the first section of this chapter, but, on the assumptions about dependence and causality outlined above, causal relationships. Arcs directed at both ends indicate dependencies that can only be explained by the presence of an unmeasured common parent variable, or, interpreted causally, the presence of a hidden common cause. (See Definition 3 above). Undirected arcs indicate that there is insufficient information to orient an arc. The algorithm also 'marks' any arcs representing dependencies that cannot be attributed to dependency on an unmeasured variable or common cause. In causal terms, these represent genuine causal relationships. (See Definition 2 above.)

The Markov Condition and the Causal Markov Condition

The Markov condition, as Richard Scheines says, is "just mathematics connecting DAGs and probability distributions" [25]. It describes the way a graph represents dependencies and independencies in a probability distribution: A variable is conditionally independent of its non-descendants, given its parents. When, on the basis of the assumptions about the relationships between dependency and causality contained in the Fundamental and Faithfulness Assumptions, a graph is interpreted causally, the *Causal Markov Condition* (*CMC*) describes the way a graph represents causal relationships: A variable is conditionally independent of its non-effects, given its direct causes. In other words, when a graph is interpreted causally, the Markov condition describes not only the representation of dependencies and independencies in data but, given the assumed relationships between dependencies and causality (and, in the case of a graph constructed using the IC algorithm, the associated definitions and the algorithm), causal relationships or causal structure.

CHAPTER 3
CARTWRIGHT'S CRITIQUE

The causal inference algorithms of Pearl and SGS rest on certain assumptions about relationships between causal structure and probabilities, including, on Cartwright's account, faithfulness and the Markov Condition. Cartwright argues not that these are never correct but that they are far less frequently correct than Pearl and SGS suppose. This chapter presents Cartwright's arguments without comment. The following chapter discusses and responds to these arguments.

The Faithfulness Assumption

In Chapter 2 we saw that the Faithfulness Assumption can be considered an application of Ockham's razor to causal inference and Pearl argues along these lines, using an analogy from perceptual inference. Suppose we see a picture of what appears to be a chair. We can take the object in the picture to be a chair (T_1) or we can regard it as two chairs aligned so one hides the other (T_2). T_1 is invariant to the angle of view, and T_2 is unlikely. In this sense T_1 is simpler than T_2. The point of the analogy is that when we observe independencies in data we have to decide whether these correspond to causal structure, that is, whether they represent an absence of causal relationship, or whether, on the contrary, what appears to be causal independence is actually an exact balancing of dependencies that conceals a causal relationship. The former, Pearl argues, is the simpler, and therefore the preferred, theory.

Another, Bayesian, argument appeals directly to the unlikelihood of the interpretation the Faithfulness Assumption rejects. For two variables *A* and *B*, for any distribution of *A*, so the argument goes, there is exactly one joint distribution of *A,B* such that the two variables are independent, but in all the remaining uncountably many distributions, the two variables are dependent. Hence, the probability of correlation, given causation, has measure one.

This argument holds if all distributions are equally likely. Nancy Cartwright, [3], however, argues convincingly that the kind of cancellation the Faithfulness Assumption precludes may occur fairly frequently.

> It is not uncommon for advocates of DAG-techniques to argue that cases of cancellation will be extremely rare, rare enough to count as non-existent. That seems to me unlikely, both in the engineered devices that are sometimes used to illustrate the techniques and in the socio-economic and medical cases to which we hope to apply the techniques. For these are cases where means are adjusted to ends and where unwanted side effects tend to be eliminated wherever possibly, either by following an explicit plan or by less systematic fiddling.

Elsewhere [4], she gives a convincing example. Birth-control pills may cause thrombosis but they also prevent pregnancy, which likewise can cause thrombosis. Depending on the strengths of the dependencies involved, these influences may cancel and, indeed, this cancellation is something we try to achieve. We try to weaken the strength with which birth-control pills cause thrombosis so that those who use the birth-control pill are no more liable to thrombosis than those who don't. In this case, "getting the cancellation that stability/faithfulness prohibits is important to us". In applied science or engineering, we may want to retain a process for some reason (perhaps it has some effects we value along with the ones dislike) or it may be easier to design a counter process to cancel certain of its effects than it is to eliminate the first process and a case can be made that the same kind of thing happens in nature. Shipley, though he remarks that such balancing occurs “only under very special conditions” [26] gives an example from nature (photosynthesis) where processes act as counterweights to establish a set point (internal CO_2 concentration). Since the theory of causal graphs cannot get far without the faithfulness assumption, this is potentially devastating.

The Markov Condition and the Factory Example

The focus of our discussion here is Nancy Cartwright’s criticism of the Markov Condition (what we have been calling the CMC) in *The Dappled World* [3], in particular, in the argument that centers on her ‘factory example’. This argument, and even this example, can be found elsewhere in her work as well. (See below.)

For Cartwright, the Markov Condition describes two kinds of conditional independencies. (She is discussing what we have been calling the *Causal Markov Condition*. This chapter follows her usage.) The first of these she justifies by the temporal relationships between causes. Causes do not operate “across temporal gaps”, so conditioning on the parents (or direct causes) of a variable makes it independent of its indirect ancestor variables (and of their descendants). The conditional independence of *B* from *A*, given *C*, depicted in Figure 1 in the previous chapter was an example of this. The second aspect is the *screening off* of variables that share a common parent: Conditioning on the common parents of two or more variables ‘screens them off’ from one other, that is, renders them conditionally independent (assuming neither sibling is a descendant of the other). Conditioning on the common parents of two or more variables, according to the Markov Condition, makes them independent. In the previous chapter, the conditional independence of *A* and *B*, given their common parent, *C* (Figure 3) is an example of this aspect of the Markov condition. It is this ‘screening off’ aspect of the Markov condition that she finds troublesome.

Screening off, she says, is ‘trivially true’ where the relationship between cause and effect is deterministic, when causes fix the values of their effects. So, for example, in the simple case of a single cause of two effects, if I know the value of the cause, I know the value of each effect. Knowing the value of the other effect tells me nothing about the first that I don’t know already, once I know the value of the cause. Hence, the effects are independent of each other, given the cause. Formally, $p(E_1 \mid C, E_2) = P(E_1 \mid C)$, where C represents the cause and E_1 and E_2 represent the effects of C. This means that the joint probability of two effects, given the cause(s) of those effects, will factor: $p(E_1, E_2 \mid C) = p(E_1 \mid C)\, p(E_2 \mid C)$. However, according to Cartwright, deterministic causality is uncommon and the Markov condition does not (in general) hold where causality is probabilistic.

When a cause operates probabilistically the relationship between the value of the cause and the value of the effect is, of course, probabilistic. For a given value of the cause, there is a probability distribution over the possible values of the effects. (In the deterministic case, there is one possible value for each effect, given the cause, with probability 1.0; any other values have probability 0.0.) Cartwright argues that the value of one effect of a probabilistic cause will almost always provide some information about the value of a second effect, even if I know the value of the cause. In other words, the effects are not conditionally independent, given their common cause; likewise, the joint (conditional) probabilities will not factor. In Cartwright's terms, there is no 'screening off'.[9]

This is so, Cartwright maintains, because of "interactions" between the "operations" whereby a cause produces its various effects.[10] In the case of deterministic causality, such interactions need not be considered since a cause, when it occurs, infallibly operates to produce all of its effects. But in the case of probabilistic causality, we have to consider not only the probabilistic relationships between a cause and each of its effects but also the relationships, likewise probabilistic, between the operations that produce the effects. When a cause can occur and yet not operate to produce one or more of its effects, such interactions have to be taken into account when calculating the probability of one or another effect, even when we know the value of the cause. The screening-off aspect of the Markov condition, according to Cartwright, assumes what she calls a "split-brain model" of the common cause", according to which a cause operates independently to produce each of its effects. However, Cartwright asserts that, in fact, "joint operations" are common. Hence, even when we know the value of the cause, knowing the value of one effect is likely to tell us something about the value of the other.

It is in this context that Cartwright presents her factory example. There are two factories, *D* and *P*. Both produce a certain chemical, *C*, that is used immediately in a sewage plant. Factory *D* produces the chemical with probability 1.0, or deterministically. Factory *P*, on the other hand, produces the same chemical with probability 0.8, that is, probabilistically. Moreover, whenever *P* produces the chemical, a pollutant, *B*, is

---

[9] In the next chapter, we explain why we regard the *factory example* as an argument against what we have called the Fundamental Assumption. We maintain that, even if correct, this argument is not fatal to causal inference algorithms.

produced as a by-product. However, the owner of factory *P* maintains that the pollutant is produced when the chemical is used in the sewage treatment plant and, in support of this claim, advances an argument that assumes the Markov condition in its screening-off aspect. The factory-owner argues that if Factory *P* did indeed produce the pollutant as well as the chemical, the probability of the pollutant would be conditionally independent of the probability of the chemical, "assuming all other causes of …. [chemical and pollutant] have already been taken into account". Thus, Factory *P* argues that if it were responsible for the pollutant then $p(C, B \mid P) = p(C \mid P)p(B \mid P)$. But in fact $p(C,B|P)$ is 0.8 while the two factors are likewise each 0.8; therefore, argues factory *P*, it is not responsible for the noxious by-product, since $0.8 \neq 0.8 \cdot 0.8$.

But factory P's argument is wrong, says Cartwright. Factory *P is* responsible for the by-product. Factory *P*'s error lies in its screening-off assumption. Because factory *P* produces both the chemical and the by-product probabilistically, knowing the value for P does not tell us whether it produced either of its two effects, whether, as Cartwright says, it "fired"; it tells us only that the probability is 0.8 in either case. But knowing the value of the by-product variable, *B*, as well tells us something about this and thereby provides information about the value of the chemical variable, *C*, that we would not know simply by knowing the value for *P*. As a result, the probabilities of *C* and *B* are not independent given *P* and hence do not factor, that is, $p(C, B \mid P) \neq p(C \mid P)p(B \mid P)$. Yet *P* is the cause of both *C* and *B*.

Cartwright buttresses this example with a negative conceptual argument in combination with a "what are the odds?" argument. First, she clears the ground with an

[10] Cartwright's use of these categories ("operations" and "interactions") presents difficulties. This, too, is

appeal to causal concepts: There is nothing in the concept of causality that implies 'screening off'; "nothing in the concept of causality, nor of probabilistic causality, constrains how nature must proceed". Then she employs an argument much like the Bayesian argument *for* Faithfulness (mentioned above) *against* "screening off". How likely is it, when we have a joint probability distribution over the effects of a common cause, that the distribution will be such that the effects are conditionally independent? The answer, she says, is "not very". She concludes, "where causes act probabilistically, screening off is not valid".

discussed in Chapter 4.

CHAPTER 4
SOME RESPONSES TO CARTWRIGHT

In Chapter 2, we discussed the Markov condition, which describes the way a graph represents independencies (and dependencies) in data. We discussed also the Causal Markov condition, which, on the basis of two assumptions about the relationships between dependencies and causality – the Fundamental Assumption and the Faithfulness Assumption - describes the way a graph represents causal relationships. Finally, we discussed the IC causal inference algorithm, which invokes both assumptions in the construction of a causal graph to which both the Markov condition and the Causal Markov conditions apply. Chapter 3 briefly outlined Cartwright's objections to causal inference algorithms. These take the form of arguments against the Causal Markov condition in its "screening off aspect" and the Faithfulness Assumption. In this chapter, we respond to Cartwright's objections.

The Factory Example and the Fundamental Assumption

In the factory example [3], factory *P* argues that if the factory were the direct causal parent of both by-product (*B*) and chemical (*C*), then, assuming the validity of the Markov Condition in its "screening off" aspect – that effects are conditionally independent, given their common cause(s) – and assuming also that "all other causes" of *C* and *B* "have already been taken into account", conditioning on the factory should make the chemical and the pollutant by-product independent. Since the chemical and the by-product are not conditionally independent, given the factory, it cannot be that the factory is responsible for the pollutant by-product as well as the chemical. Yet, Cartwright informs us, the factory *is* responsible for the by-product. Factory P's argument is faulty because its "screening off" assumption is not valid where causes act probabilistically, as they do here and, in fact, as they do in most instances. Because factory P produces both the chemical and the by-product probabilistically, there is a probabilistic relationship between the operations whereby it produces each. "Knowing that the cause occurred," that is, that factory P was operating on a given day, "will not tell us whether the product [the chemical] resulted or not. Information about the presence of the by-product will be relevant since this information will tell us (in part) whether, on a given occasion, the cause actually 'fired'".

That information about the presence of the by-product is not sufficient of itself to establish whether factory *P* "fired" and produced the chemical, suggests that the by-product may be absent in the presence of the chemical or present in its absence or that one or both may have a source other than factory *P*. The latter possibility is consistent with the assumption that any other causes of *C* and *B* have been taken into account, since that assumption does not exclude other causes. More importantly, if we take that assumption to mean that any other causes of *C* and *B* are included in the conditioning set with factory *P*, we can conclude that the chemical and the pollutant variables are dependent in all contexts and step 1 of the algorithm would place (or, more precisely, leave) an arc between the chemical and the by-product variables. The graph would look like the one in Figure 4.1.

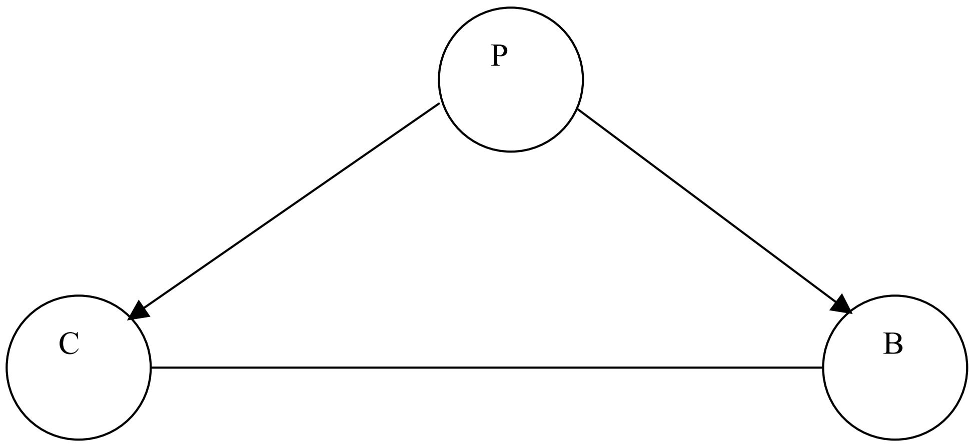

Figure 4.1: The Chemical and the By-product are dependent in all contexts

From the perspective of this graph, it is far from obvious that the factory example is a counter-example to the Markov condition. The independencies and dependencies represented in the graph, according to the Markov condition, are consistent with those of the example; in particular, the chemical and the pollutant are not conditionally independent, given the factory. In other words, the graph does not imply what Cartwright terms "screening off". Accordingly, our initial response to the factory example argument was that the example does not challenge the Markov condition. If the chemical and the by-product are not conditionally independent, the algorithm will place an arc between them.[11] Factory *P*'s argument assumes that effects of a common cause must be conditionally independent, given their common cause, ignoring the possibility of a causal

[11] The IC algorithm would not orient any of the arcs but we have oriented the arcs between factory *P* and the *Chemical* and between *P* and the *By-product* to accord with the causal information given in the example.

relationship between the effects or of another, hidden, common cause (though Cartwright seems to have ruled out the latter possibility in this example).

However, a later paper [2] clarifies the example somewhat. There, Cartwright provides a graph of the "true" (the quotations are hers) causal structure. (See Figure 4.2) In this graph there is no arc between chemical and by-product, which means that, according to the Markov condition, they *should* be conditionally independent, given the factory.[12] It would appear, then, that her argument is made with respect to the "true" causal structure. It is the Markov condition as applied to the "true" causal structure of the example that implies the conditional independence, given the factory, of chemical and by-product, contrary to the statistical dependencies given in the example and (correctly) represented in Figure 4.1. The discrepancy between the "true" causal graph and the graph representing the dependencies and independencies of the example suggests another interpretation of Cartwright's argument.

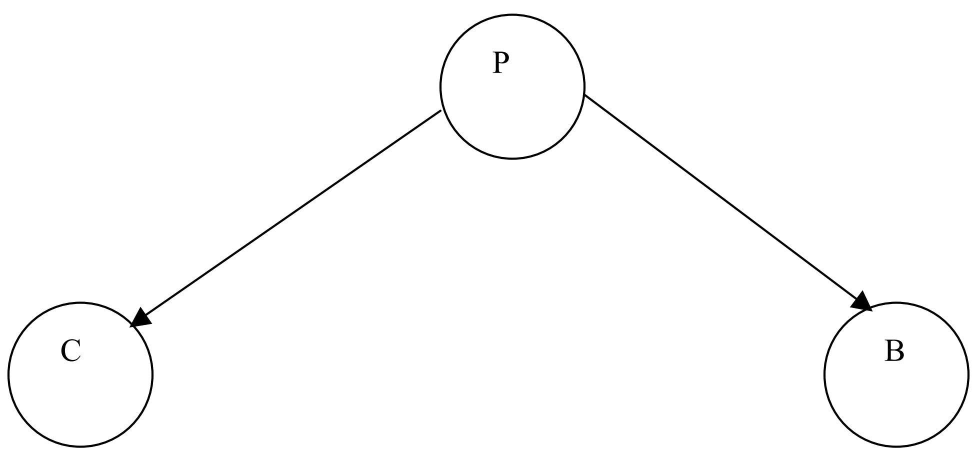

Figure 4.2: Cartwright's "true" causal graph

[12] Cartwright's graph includes the other factory (Factory *D*), but for our purposes, that is unnecessary.

On this interpretation, Cartwright has no quarrel with the Markov Condition, which describes a graphical representation of dependencies and independencies in data. Her quarrel is with the Causal Markov Condition, which says that the kind of graphical representation the Markov Condition describes will also represent causal relationships.[13] The latter describes a "matching" between causal structure and dependency graph that, according to Cartwright, does not exist in the factory example. In the factory example, the chemical and by-product variables must be conditionally independent, given the factory, if the dependency graph is to correspond to the "true" causal structure. Because this is not so, the Causal Markov condition is invalid.

The difference is located in the arc representing the dependency between the chemical and the by-product. Since this arc is present in the graph that represents dependencies and independencies between variables but absent in the graph representing "true" causal relationships, and since the example stipulates that there is no hidden common cause, this arc must correspond to a non-causal dependency. The argument thus seems to be that what makes the Causal Markov condition invalid is the presence of a dependency that exists in all contexts and does not reflect or derive from a causal relationship between the chemical and the by-product or from a hidden common cause of both. This amounts to a denial of what, in chapter one, we called the "Fundamental Assumption". We shall return to this later.

It seems reasonable to associate this non-causal dependency with the probabilistic relationship that, according to Cartwright, exists between the operations of the factory to produce the chemical and the by-product. The factory example is presented as an example of a cause that operates probabilistically to produce its effects. Cartwright maintains that, where causality is probabilistic – as, she claims, it usually is – there is a probabilistic relationship between the operations whereby it produces its various effects and that it is because of such "joint operations" that the (Causal) Markov condition, in its "screening off" aspect, is not valid.

A subsequent article [4] is consistent with this interpretation and further elucidates Cartwright's position. Having critiqued the faithfulness assumption and the assumption that causality implies dependency that follows from it, she turns to the question of whether dependency implies causality - what we have termed the "Fundamental Assumption" – and the Causal Markov condition.[14] She begins with the uncontroversial point that a probabilistic dependency between *A* and *B* may derive from different circumstances. It may be that *A* causes *B* or vice versa but there are other possibilities. Of those she enumerates – and these are not uncontroversial – two have a particular bearing on the factory example. One is that *A* and *B* have a common cause "where either the causes are deterministic or the action of producing *B* is independent of the action of producing A". The other is that *A* and *B* are "produced as product and by-product from a probabilistic cause".

Cartwright contends that "advocates of Bayes nets acknowledge and try to deal … squarely" only with the first. The Causal Markov condition, which states that a variable is conditionally independent of all variables except its own effects, given its

[13] Cartwright doesn't distinguish the Markov condition from the Causal Markov condition and speaks throughout of the "Markov condition" but it seems that what she has in mind is what we have been calling the "Causal Markov condition".
[14] It is perhaps significant that the later paper does use this terminology.

direct causes, discriminates dependencies deriving from a causal relationship from dependencies that derive from a common cause under the conditions given (deterministic causality or independent operations of a single cause). If a dependency derives from a common cause (or causes) of this kind, and not from a causal relationship between the dependent variables, then the dependent variables will be conditionally independent, given the common cause(s). Dependencies that correspond to causal relationships between the dependent variables are those that persist when causal parents are conditioned on.[15] "To claim that this is enough to ensure a causal connection," she says, "is to maintain the causal Markov condition".[16]

Cartwright claims that it is not enough. The Causal Markov condition, she argues, holds under the stated conditions. However, there are dependencies that survive all conditioning, dependencies that exist in all contexts, which nonetheless do not derive from a causal relationship. Where, for example, the dependent variables are the product and by-product of a probabilistic cause, the product and by-product will *always* be dependent, even when their common cause is conditioned on. Contrary to the Causal Markov condition, these dependencies do not derive from a causal relationship. This is precisely the situation Cartwright describes in her factory example.

In an earlier work [1] she discusses "joint operations" of another kind and a resulting non-causal dependency.

> "A typical case occurs when a cause operates subject to constraint, so that its operation to produce one effect is not independent of its operation to produce another. For example, an individual has $10 to spend on

[15] The Causal Markov condition might also be considered to discriminate dependencies deriving from an indirect causal relationship from those deriving from a direct causal relationship. This is suggested by Cartwright's discussion of the first aspect of the Markov condition in *The Dappled World*.
[16] "Advocates of Bayes nets" would say that a hidden common cause may also be a possibility in such cases, but Cartwright is not concerned with that possibility here.

> groceries, to be divided between meat and vegetables. The amount that he spends on meat may be a purely probabilistic consequence of his state on entering the supermarket; so too may be the amount spent on vegetables. But the two effects are not produced independently. The cause operates to produce an expenditure of *n* dollars on meat if and only if it operates to produce an expenditure of 10 – *n* dollars on vegetables. Other constraints may impose different degrees of correlation.

In this example, the common cause is the shopper's state of mind on entering the store while the (non-causally dependent) effects are the amount spent on meat and the amount spent on vegetables. These are given as *a* = *the amount spent on meat* and 10 – *a* = *the amount spent on vegetables* but could just as easily have been *a* = *the amount spent on vegetables* and 10 – *a* = *the amount spent on meat*. The relationship between the operations that produce these effects is described as an "if and only if" relationship. The value of each effect implies the other; they are perfectly inversely correlated measurements. Hence, each effect is conditionally independent of the only other variable involved, the common cause of both (the shopper's "state"), given the other. According to the Markov condition, then, when the example is represented graphically, each is parent of the other and the result is an isolated cycle, a cycle between two variables unconnected to any other variable.[17] Recall that in Chapter 2, we mentioned that the IC algorithm assumes that there are no directed cycles and that this assumption corresponds to an intuition that a thing does not cause itself. These observations suggest, first, that the example is ruled out as an appropriate subject for the IC algorithm and, secondly, that the two effect variables in fact represent different ways of measuring the same thing and are properly considered a single variable, in other words, that they are proxies for the same entity.

[17] The IC algorithm would eliminate any arcs other than the one between the two effect variables but would not orient it. (See Chapter 2.)

Likewise, in the factory example, if the by-product is produced if and only if the chemical is produced, it is arguable both that the example is excluded a priori as appropriate material for the IC algorithm and that the variables *Chemical* and *By-product* are proxies for the same thing. However, the example is unclear on this point. In [5] Cartwright says that the chemical and by-product are "always produced together" and in [3] she says that the by-product is produced whenever the chemical is produced and we might expect that the by-product is not produced unless the chemical – of which it is a by-product – is produced. On the other hand, Cartwright appears to say at one point that knowing the value of the by-product variable is not enough to establish the value of the chemical variable. (See above.) It may be significant that many, if not all, Cartwright's examples of non-causal dependencies seem to involve if and only if relationships between the dependent variables (though in most cases she says that some less than perfect dependency would serve as well to make her point) since all such examples are open to these same objections.

Cartwright's response to an argument similar to the "single variable" objection above, made with respect to another example involving perfectly correlated side effects, is that while this would save the Causal Markov condition, it would not allow us to study "causal relations among the kinds of quantities we are generally interested in" [5]. Indeed, we may be interested in these quantities, but it does not follow that they are appropriately investigated by the kinds of methods under discussion, which require variables defined in the way chapter two describes.

Before we leave the question of variables, we should make some mention of a common technique for dealing with the kind of difficulty Cartwright presents in her factory example [3]. According to the story, factory *P* may or may not 'fire' and produce the chemical (and likewise, the pollutants). Is the occurrence or non-occurrence of this event adequately represented by the *Chemical* (*C*) and *By-product* (*B*) variables, representing, respectively, the production or presence of the chemical and the by-product? As we saw earlier, Cartwright says that "information about the presence of the by-product will be relevant since this information will tell us (in part) whether, on a given occasion, the cause actually 'fired'". That being the case, and since the 'firing' of factory *P* is an event that may or may not occur, perhaps it should be represented by a *Fired* (*F*) variable, as in Figure 4.3. P(*fired* | factory *P*) would then be 0.8. (It is probably not necessary to have a 'fires' variable for the other factory.) The *Fired* variable would determine (or fix) the value of its effects. The probability of each would be 1.0, given *fired = 1.0*. This is the kind of situation where, Cartwright avers, "screening off" is "trivially true". Once we knew whether factory *P* had 'fired' we would know the values of both the chemical and by-product variables. Knowing the value of one of the effects would therefore provide no further information about the value of the other. In the case of the alternative scenario, where the by-product does not invariably accompany the chemical, the byproduct would be produced with some probability less than 1.0 but it would still be true that, once the value of the "fired" variable was known, there would be nothing further to be known about the value of the *Chemical* variable from the value of the *By-product* variable and vice versa. The *Fired* variable would be a common cause of both the chemical and the pollutant that would render them conditionally independent when conditioned on. This is a "move" made by Hausman and Woodward in an early response to Cartwright's factory example [13].[18] Thus, Hausman and Woodward reify

[18] They have lately taken a different tack that is beyond the scope of the present work.

Cartwright's "processes" and "interactions" – we shall say more about these in what follows – and construct a complete explanation of the observed data consistent with the Bayes Net formalism, but also explicating and answering Cartwright's domain-specific concerns. However, Cartwright might reply that there is no way to determine directly whether the factory has 'fired' apart from the presence of the chemical or the by-product.

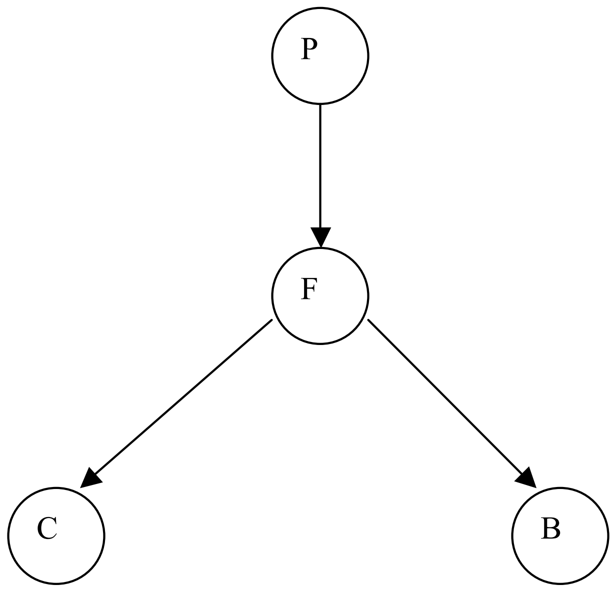

Figure 4.3: *C* and *B* are conditionally independent, given *F*

Here is another variation of the same "move": Suppose the arc between *Chemical* and *By-product* is interpreted as representing a potential cause. A potential cause may be either a genuine cause or a hidden cause. We have suggested that it might represent a genuine causal link but suppose it represents a hidden common cause. Then the "true" causal model might be as shown in Figure 4.4.

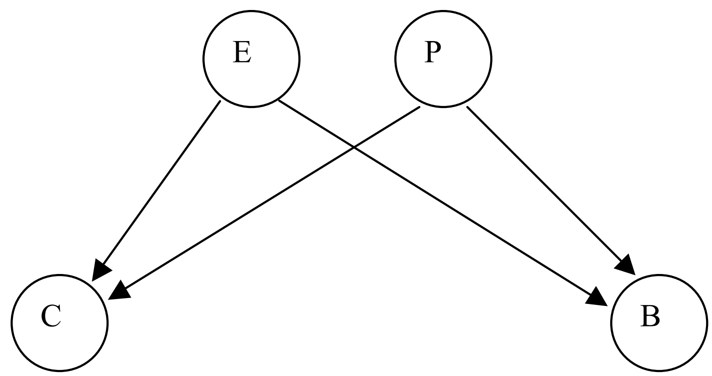

Figure 4.4: *C* is conditionally independent of *B*, given *E* and *P*

Now, the distribution of each of *Chemical* and *By-product* must be specified in terms of the joint distribution of something like barometric pressure (*E*, on the supposition that this kind of *Environmental* factor has an influence) and factory *P*. This is less vulnerable to the charge of employing a handy recipe that can be used to make any two dependent variables independent than is Hausman and Woodward's solution, since it explicates a possibility inherent in the definition of potential cause. It is true that Cartwright stipulates against hidden causes of *C* and *B*. However, it might be said that the "interactions" between "operations" that are so prominent in her discussion of the factory example look suspiciously like hidden causes.

This brings us to what we might describe as our second response to Cartwright. Whereas the causal graph world consists of arcs and variables and *CPT*s (*Conditional Probability Tables*), Cartwright introduces "operations", "probabilistic relations" between "operations", and "joint operations", which somewhat resemble variables and dependencies, without making precise the nature of these new entities, their relationships to each other, or their differences from, and relationships to, the old categories. How do we distinguish an operation from a variable (including a latent variable) or a dependency (correlation)? How do we distinguish a probabilistic relationship between operations from a dependency (correlation) between variables? Can we measure them and can we represent them in such a way as to eliminate potential graphs from consideration as viable causal models? The problem, in our view, is that Cartwright has expanded the ontology without precisely defining the categories she introduces. The Causal Markov condition is an integral component of a formally expressed set of definitions and algorithms, yet she provides nothing comparable by way of formalism for her own categories, though they play a crucial role in her critique of the Causal Markov condition and, by extension, the definitions and causal inference algorithms with which it is associated. This makes it difficult to interpret or respond to her claim that the existence of such entities or processes fatally undermines the claims made on behalf of causal inference algorithms.

A related point is that, although the "operations" of a cause and the "probabilistic relations" between those "operations" are central to Cartwright's argument, neither is represented in causal graphs, or so it would appear from her factory example. Again, what are they? Can we measure them? If they are as devastating to causal inference algorithms as she maintains and if they can be measured, perhaps the solution is to represent them in causal graphs, altering or expanding the algorithm as necessary. Even if her arguments against the Causal Markov condition are valid, whether the Causal Markov condition would hold for these expanded causal graphs is an open question. But neither this question nor the prior question of the validity of her criticisms of causal inference algorithms and causal graphs as they now stand can be addressed in the absence of precise definitions for the expanded ontology she proposes.

There is also the question of how Cartwright understands causality. The claim that the dependency between the chemical and the by-product is non-causal is at the heart of her argument. Cartwright tells us that even when we know the value of the factory variable (and any other causes of the chemical and the by-product), knowing the value of the by-product variable tells us something more about the value of the chemical variable. Are they causally related? Cartwright says that they are not.

Yet, Cartwright insists that causality as multifarious [3]:

> Causes make their effects happen. That is more than, and different from, mere association. But it need not be one single different thing. One factor can contribute to the production or prevention of another in a great variety of ways. There are standing conditions, auxiliary conditions, precipitating conditions, agents, interventions, contraventions, modifications, contributory factors, enhancements, inhibitions, factors that raise the number of effects, factors that only raise the level, etc.

In view of this, we question whether the relationship between chemical and pollutant in the factory example really is a non-causal relationship.[19] We further question whether and to what extent it matters, for our purposes. At the very least the dependency between them is worth knowing. It contributes to predicting and explaining. It helps us to navigate reality (or it would if the factory example were real, as opposed to imaginary). That, arguably, is the purpose of causal graphs and therefore the arc that represents that dependency is not something we would want to leave out.

But perhaps there are dependencies that persist in all contexts and that are not causal in any sense and that, if interpreted as indicating a causal relationship or a hidden common cause, would mislead us. Perhaps, in other words, there are exceptions to the

Fundamental Assumption of a sort that would lead to significant errors in the output of the kind of causal inference algorithms we are discussing. This brings us to our primary response to Cartwright.[20] Even if exceptions to the Fundamental Assumption, and therefore to the Causal Markov condition, exist, the IC algorithm and causal graphs are nonetheless valuable, though not infallible, tools in the process of causal discovery.

In fact, it is not easy to imagine how human beings (or, indeed, other animals) would live without something like the Fundamental Assumption. If, as Cartwright argues, it not without exception, it is nevertheless essential. Hausman and Woodward [14] make a similar point when they argue that "even though fallible…, the causal Markov condition may still be our best hope for making causal inferences from non-experimental data, at least when strong domain-specific background knowledge is lacking". We would add that this is precisely the setting in which the algorithms are intended to be deployed. Where causal structure of the world already is well known, there is no need for causal inference algorithms. These are useful when we lack knowledge of true causal structure apart from probability distributions.[21]

Causal discovery, as performed by human beings, is an iterative process. We should not be surprised, then, if causal discovery as performed by causal inference algorithms likewise requires iterative correction and improvement. We maintain that causal inference algorithms can assist, but not replace, human knowledge and reasoning. Human beings drive the iterative process, selecting and preparing data, interpreting

---

[19] This issue is complicated by the fact that, as we have said, it is far from obvious what that relationship is.
[20] For the moment, we ignore the difficulties inherent in Cartwright's account of such non-causal dependencies in terms of an expanded and non-specific ontology and likewise disregard the particularities of the factory example.
[21] However, it must be admitted that if exceptions to the Causal Markov condition are as common as Cartwright claims, their utility would be correspondingly limited. On the other hand, this question is all tied up with the questions about her terminology discussed above.

results, detecting possible errors, formulating hypotheses, testing, and initiating the next iteration. Human insight is particularly useful in locating problems and overcoming limitations deriving from exceptions to general rules and assumptions. In sum, exceptions to the Fundamental Assumption, and therefore to the Causal Markov condition, would not be fatal to the IC algorithm (or to similar algorithms).

Of course, if such exceptions exist, it would be useful to understand their nature. This is where our secondary response to Cartwright comes in. Her use of new (for this area of investigation) and undefined categories makes it is difficult to grasp the nature of the exceptions for which she argues and it is correspondingly difficult to assess their existence or prevalence.

The second prong of Cartwright's critique is an argument that exceptions to the Faithfulness Assumption are not uncommon. Our response to this argument parallels our primary response above and includes an example of human intervention in the computer-assisted process of causal discovery.

The Faithfulness Assumption

Cartwright argues that "cancellations" incompatible with the Faithfulness Assumption are frequent. This argument is clear and plausible and we accept that the sort of cancellations she describes may indeed occur frequently. It may be easier in applied science and/or engineering to design a counter process to cancel certain effects of a process than it is to eliminate the process causing the effect in the first place. Thus, a case can be made that Nature, as engineer, frequently uses the same ploy. Yet we do not regard the existence of such exceptions to Faithfulness as fatal to causal inference algorithms anymore than we do the possible existence of exceptions to the Fundamental Assumption and for the same kinds of reasons. As we argued with respect to the

Fundamental Assumption, something like the Faithfulness Assumption plays an important role in human causal inference and it is difficult to see how we – or a causal inference algorithm - could do without it. Further, and as we have also said, causal inference algorithms, like human causal inference, are iterative. Finally, they should be regarded as useful tools rather than as replacements for human reasoning. They require human co-operation and assistance, particularly where exceptions to general rules or assumptions are concerned.

Here is an example of how this can work. In this example, the user employs a visualization tool developed by Neufeld and Guan [12]. Dennis *et al* [9] report a meta-study of studies of the relationship between sunscreen use and melanoma. While it has been assumed that sunscreen use reduces the risk of melanoma, some reports suggest that sunscreen use actually increases the risk for melanoma. However, the meta-study determined that several such reports failed to take into account different sunscreen usage rates as between light-skinned and dark-skinned people, who are more and less susceptible to melanoma, respectively. Figure 4.5 shows the qualitative causal relationships involved between sunscreen use *(S)*, melanoma *(M)* and darkness of skin *(D),* but does not indicate the sign or magnitude of the causal relationships. In Cartwright's terms, Figure 4.5 shows the "true" (as above, the quotation marks are hers) causal structure.

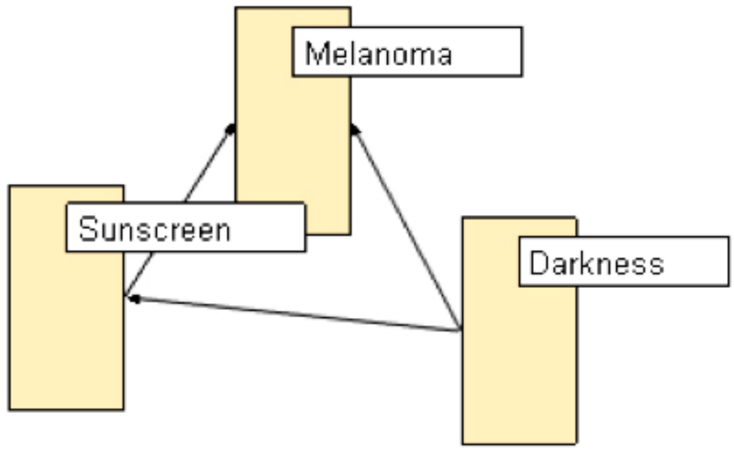

Figure 4.5: Qualitative causal model of sunscreen, melanoma and darkness of skin

Our example supposes that reports showed *no* relationship between sunscreen use and incidence of melanoma. (This is almost as surprising as the actual reports.) Our supposition is that higher melanoma rates amongst those with higher sunscreen usage rates (light-skinned people) exactly balance lower melanoma rates amongst those with lower sunscreen usage rates (dark-skinned people), rendering sunscreen usage and melanoma independent. Thus recast, it is an instance of the kind of violation of the Faithfulness Assumption advanced by Cartwright, that is, it is an example of “canceling” or “balancing” dependencies along different causal pathways and resulting independence between causally related variables.

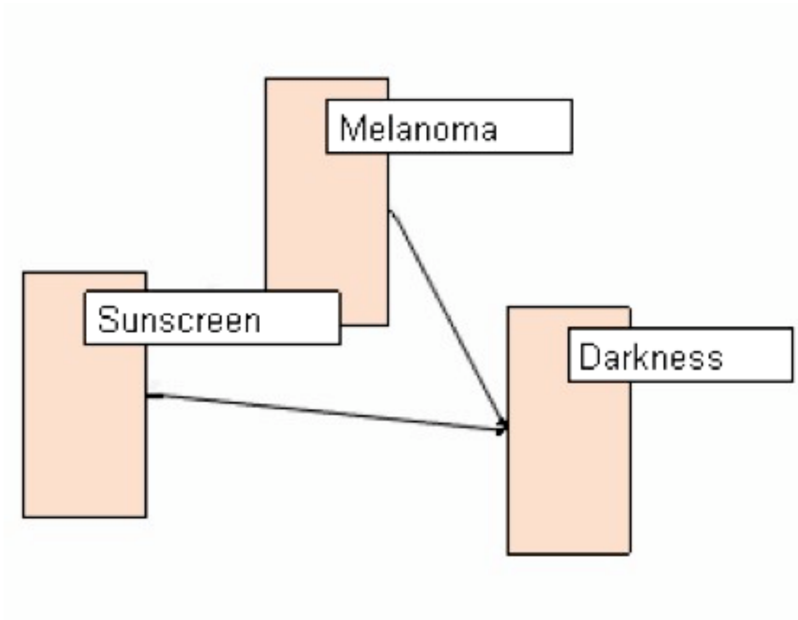

Figure 4.6: Causal model output by IC algorithm

The causal model output by the IC algorithm would then look like the one in Figure 4.6. The model shows a world where *melanoma* and *sunscreen* are independent common potential causes of *darkness* of skin, and mathematically is a minimal representation of possible causal relationships in a world coherent with the raw data. (Chapter 2 includes a simplified description of the reasoning involved.) Because the example violates the Faithfulness Assumption, this model is incorrect. To a human being, this causal configuration will appear highly implausible.[22] Accordingly, using the visualization tool mentioned above, the user constructs what he or she suspects may be

[22] In domains where the variables are not so well understood, the user can use the model to see what predictions it makes about interventions, and test these predictions in the actual domain. This would involve experimenting with *setting*, which we discuss below.

the correct causal model in order to explore the relationship between *Sunscreen* and *Melanoma.*

However, before we describe this, we need to introduce a distinction between *seeing* and *setting*. Seeing is the common procedure of computing a conditional expectation—given a subpopulation of a known population, what is the posterior distribution of all related variables in the subpopulation? For example, we may wish to compute p(M|see(S)), the probability that someone we see using sunscreen might develop melanoma. This can be computed as the ordinary conditional probability *P(M|S)*:[23]

$$p\big(M \mid see(S)\big) = p(M \mid DS)\,p(D \mid S) + p(M \mid \neg DS)\,p(\neg D \mid S) \qquad (4.1)$$

In this discrete formulation, *D* indicates dark skin and $\neg D$ indicates light skin.

Setting is about the consequences of actions (or interventions). For example, we may wish to find out whether sunscreen use reduces the risk of melanoma. Contrary to what might be expected, this is not the same as finding the probability of melanoma amongst sunscreen users. We alluded to the difference above when we said that the probability of melanoma for sunscreen users does not indicate whether sunscreen use reduces the risk of developing melanoma since light-skinned people, who are at greater risk for melanoma, are more likely to use sunscreen than are people with dark skin, for whom the risk is smaller. Although $p(M \mid DS)$ correctly counts the incidence of melanoma among dark-skinned sunscreen users, multiplying it by $p(D \mid S)$ will not predict the posterior incidence of melanoma among dark-skinned sunscreen users in the new (*setting*) scenario, wherein everyone uses sunscreen, since it does not count dark-skinned persons currently not using sunscreen. Thus, $p(M \mid DS)$ should be multiplied by

[23] For this example, we use discrete probabilities, but the argument generalizes to continuous variables.

the proportion of all persons who are dark-skinned, assuming that the effect of *Sunscreen* among all the *D*s will have the same effect as it does among those *D*s who currently use sunscreen. A symmetric argument applies to the second multiplicative term. Hence, to compute the effect of setting (as opposed to observing or seeing) *S*, we calculate:

$$p(M \mid see(S)) = p(M \mid DS)p(D) + p(M \mid \neg DS)p(\neg D) \quad (4.2)$$

Pearl [23] shows that in terms of graphical structure, this is equivalent to removing the arc from *Darkness* to *Sunscreen* and recomputing the conditional probabilities accordingly. When we set *Sunscreen*, sunscreen use does not depend on skin colour. Thus, we are able to quantify the consequences of an intervention on the basis of observational data.

Now we can return to our user's exploration of the relationship between *Sunscreen* and *Melanoma*, using what she suspects is the correct model. Figure 4.7 shows the results of her first explorations with *seeing*. That the user has fixed the value of the *Sunscreen* variable is indicated by the green colour of the *Sunscreen* rectangle and by a line and a numeral showing the user-determined value for the variable. The other rectangles show the distribution of the *Melanoma* and *Darkness* variables given the value of the *Sunscreen* variable. The user grabs the value of the *Sunscreen* variable, drags it up and down, and finds from the contrived data that *Sunscreen* and *Melanoma* are unrelated. Changes to *Sunscreen* change *Darkness* significantly – the bar goes up when *Sunscreen* goes down and down when *Sunscreen* goes up - but the bar in the *Melanoma* rectangle barely moves.

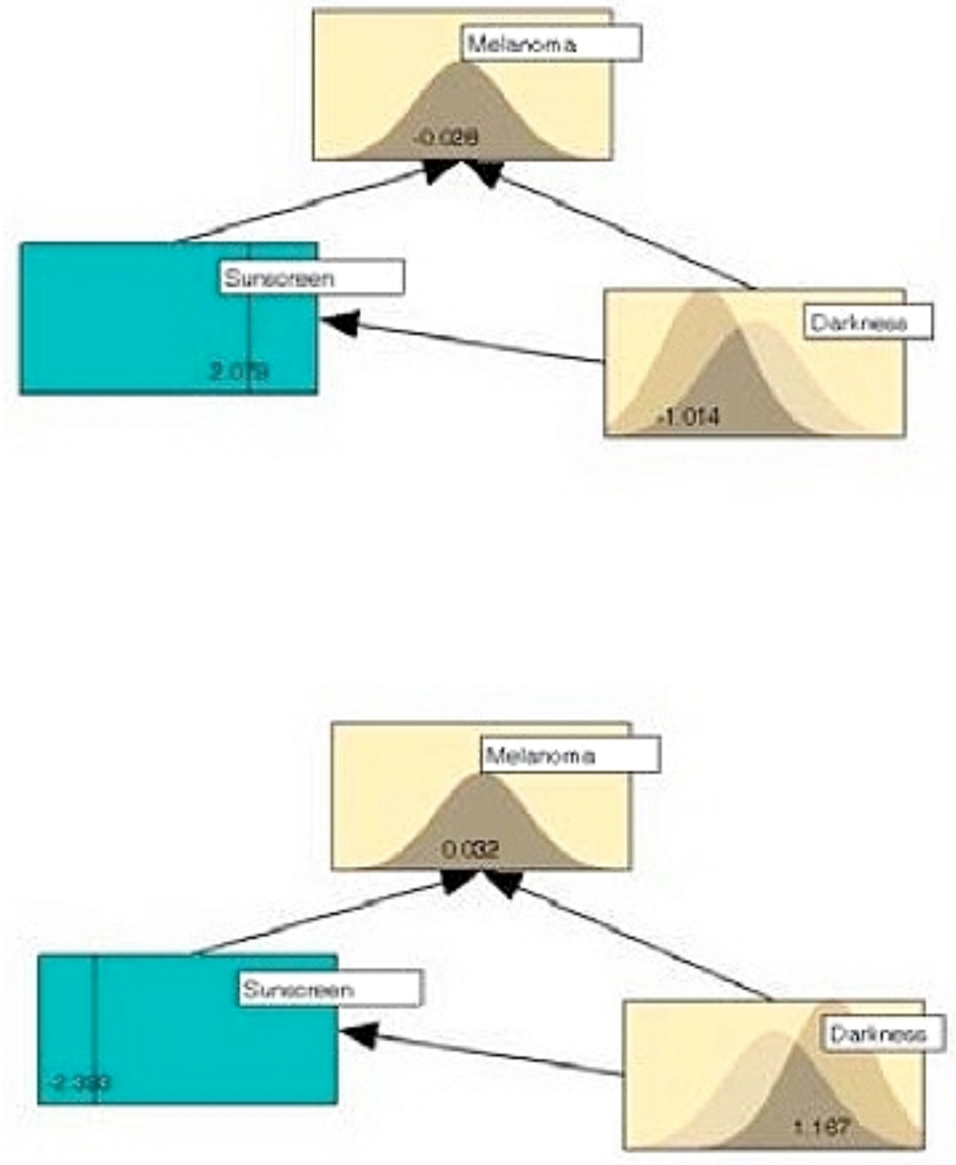

Figure 4.7: *Melanoma* doesn't respond much to *Sunscreen*

The statistical explanation, as we have said, is that *Darkness* acts as a confound. That is, light-skinned individuals are more likely to use sunscreen, but they are also more likely to develop melanoma than dark-skinned people. An experienced data analyst would ask what happens when *Sunscreen* is manipulated after *Darkness* is first fixed at some value. This is shown in Figure 4.8.

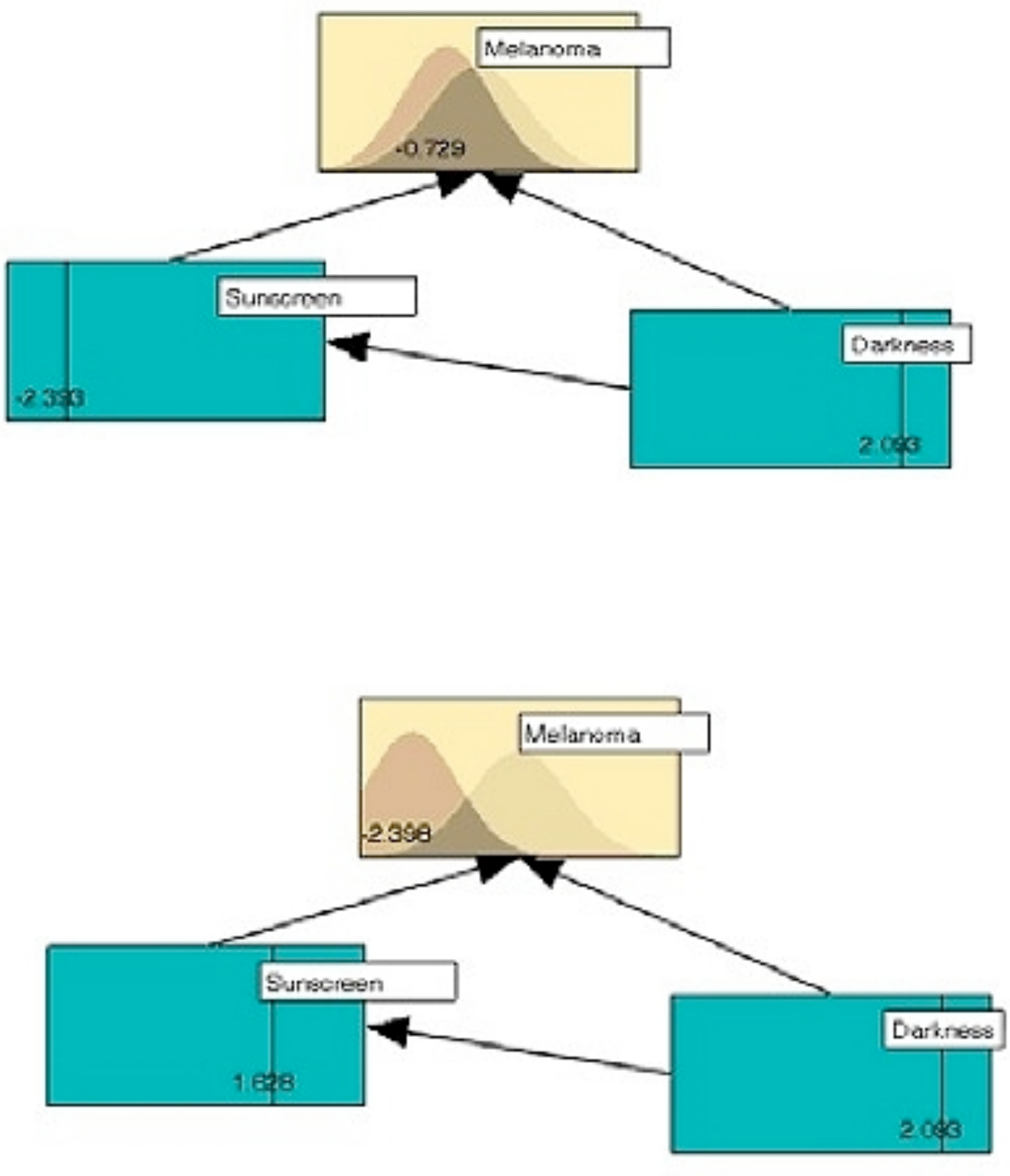

Figure 4.8: Manipulating *Sunscreen* after fixing *Darkness*

The user must first choose to *see* a fixed value for the *Darkness* variable, and then *see* a range of values for *Sunscreen by* dragging it up and down. The before and after images in Figure 4.8 now reveal the relationship between *Sunscreen* and *Melanoma*. Whether *Darkness* is high or low, fixing *Darkness,* and then increasing *Sunscreen* results in a decrease to *Melanoma.* (The "shadow distribution" shows the distribution of the variable in the absence of such "fixing", that is, it shows the unconditional or prior distribution of the variable represented.) Manipulating a variable by *seeing* a value corresponds to computing conditional values of all other variables in the domain, given the observed value. So when the user fixes *Darkness*, expected values for *Sunscreen* and *Melanoma* are calculated based on the selected value for *Darkness*. When the user then sees different values for *Sunscreen, Melanoma* responds as expected. Moreover, the same result will obtain for any values the user chooses for the *Darkness* variable.

The reason the user needs to perform a double seeing operation is that there are two paths of probabilistic influence from *Sunscreen* to *Melanoma* that cancel each other out. One is the direct path, and the other is the indirect path from *Sunscreen* to *Darkness* to *Melanoma*. (This is all the more confusing because this probabilistic path goes "against the arrows".) By fixing the *Darkness* variable, the positive probabilistic path of influence from *Sunscreen* to *Melanoma* via *Darkness* is effectively blocked, and the software correctly shows this relationship.

Because there are few potential confounds in this three node world, trying all *see* operations is not logistically difficult. However, in a richer dataset, this process may be cumbersome. This is where *setting* becomes important. If the user simply *sets* the *Sunscreen* variable, *Melanoma* responds appropriately, and more importantly, consistently with the cause-effect relationship that actually exists in this (contrived) world. The before/after graphs in Figure 4.9 show the visualization.

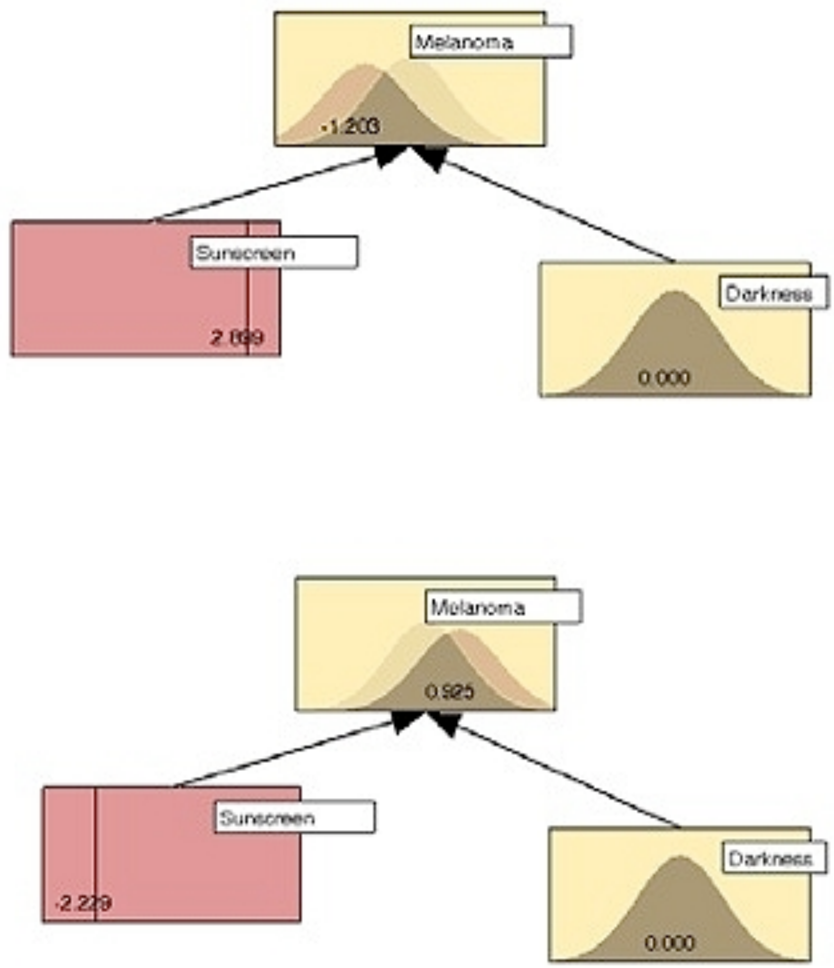

Figure 4.9: Results of *setting Sunscreen*

To understand why this happens, revisit the semantics of setting as given above. The definition of *setting* cannot be derived mechanically from the axioms of probability, but rather uses probabilities from a known world in a novel way to make predictions about an unknown world resulting from an intervention. In terms of probabilistic influences, *setting* accomplishes the same thing as the double *seeing* operation by canceling the positive probabilistic influence *Sunscreen* has via the path through *Darkness*. We mentioned earlier that this method has a simple and elegant interpretation in terms of the graphical models—the arc between *Darkness* and *Sunscreen* is erased, explicitly eliminating the positive causal path.

The user is now in a position to understand not only the true causal relationship between sunscreen and melanoma but the role of skin colour as well. She can correct the output – and the input for the next iteration, if there is one, of the algorithm.[24]

[24] Such human correction does not depend on the visualization software to which we have referred but is greatly facilitated by it.

CHAPTER 5
CONCLUSIONS

Cartwright does not appear to object to the Markov condition, which describes a way of representing dependencies in a graph. Rather, she objects to what is commonly referred to as the Causal Markov condition, which describes the way a dependency graph represents causal relationships. Causal interpretation of a dependency graph depends on two assumptions about the relationship between dependency and causality (together with some definitions) and her arguments are directed against those two assumptions: what we have termed the Fundamental Assumption and the Faithfulness Assumption. According to the first, dependency implies causation, that is, if A and B are dependent, then A causes B, B causes A, or A and B have a hidden common cause.[25] According to the second, all independencies reflect causal structure.

The Factory example is presented as an example of dependencies that do not correspond to causal relationships in the way the Causal Markov condition describes. Some exegetical research led us to interpret it as an example, or at least an intended example, of dependency incompatible with the Fundamental Assumption, that is, dependency in the absence of causality. More precisely, the dependency between the variables representing the product and by-product of the factory is presented as existing in all contexts (that is, when the factory is conditioned on) yet non-causal – according to the example, there is no causal relationship between the product and by-product, nor is there a hidden common cause. Cartwright claims that, where a cause operates probabilistically rather than deterministically, as, she asserts, most causes do, there is a probabilistic relationship between the operations whereby it produces its various effects. It is because of such "joint operations" that the effects of a probabilistic cause are dependent in all contexts even where there is no causal relationship between the effects.

Causal graph algorithms pertain to variables defined as described in Chapter 2 and presume that there are no directed cycles between variables. There are good reasons to think that many of Cartwright's examples of non-causal dependencies fail to meet one or both requirements; both Cartwright's factory example and her omnivorous shopper example are arguably vulnerable to this kind of objection.

That aside, we have two main responses to Cartwright's argument. Our secondary response addresses the critical dependence of her critique of the Causal Markov condition on a set of categories ("operation", "joint operations", "probabilistic relations") exogenous to the formally defined structure of definitions and algorithms of which the Causal Markov condition is a part and against which her critique is ultimately directed. This extension of the ontology, combined with a lack of formal definitions for the new categories, makes her argument difficult to assess.

A related point can be made about her understanding of causality. It is certainly no argument against Cartwright that she provides no definition of causality! However, she insists, on the one hand, and not unreasonably, that causality is polymorphic while, on the other, she asserts that the dependency between the product and by product of the exemplary factory is non-causal and, more generally, that effects of a probabilistic

[25] In fact, this is the informal or loose version of that assumption. Strictly speaking, the dependency is dependency in all contexts and the causal relationship is direct, as between measured variables.

common cause are non-causally dependent. In light of the first claim, the second would seem to require some justification or explanation beyond what is provided.

Our primary response is that, even if there are exceptions to the Fundamental Assumption, and therefore to the Causal Markov condition, neither these nor the causal inference algorithms that depend on them are thereby rendered useless. It is difficult to see how human causal inference, or, therefore, human life, could proceed without the Fundamental Assumption and the possibility that it might not be valid without exception is not particularly troubling. Few of the assumptions that are crucial to human reasoning meet such an exacting standard and it would not be altogether surprising if the Fundamental Assumption proved to be one of many such indispensable but not infallible principles.[26]

If assumptions that are not universally valid are necessary and legitimate in human causal inference, we can expect them to be similarly necessary and legitimate in causal inference algorithms. While it is true that computers are singularly and notoriously ill equipped to deal with exceptions to general rules, causal inference algorithms can assist human reasoning about causality, though they cannot replace it. They are, in other words, tools, meant to be used by human beings as part of an iterative process of causal discovery. If there are exceptions to the Fundamental Assumption, then this is an area where human contributions and interventions will be required if the tool is to achieve its results, where such exceptions exist. But this does not mean that the tool is not eminently useful for all that: The current algorithms provide ways to mechanically process large quantities of data and construct a causal models capable of predicting the results of

[26] Of course it would be eminently useful to understand the nature of such exceptions and the frequency with which they occur and, as we noted in the previous chapter, our secondary response is relevant here.

interventions in a meaningful way. This is a marvelous achievement, even if falls short of replacing human thought.

Our response to Cartwright's convincing argument against the Faithfulness Assumption parallels this primary response to her critique of the Fundamental Assumption, augmented by an example of how human intervention can correct and interpret the output of the algorithm where a violation of Faithfulness leads to an incorrect result. At the same time, the example demonstrates the capacity of a model to predict the consequences of an intervention.

REFERENCES

[1] Nancy Cartwright. Nature's Capacities and their Measurement. Oxford University Press, 1989.

[2] Nancy Cartwright. Causal Diversity and the Markov Condition. *Synthése* 121: 3-27, 1999.

[3] Nancy Cartwright. *The Dappled World: A Study in the Boundaries of Science*. Cambridge University Press, 1999.

[4] Nancy Cartwright. What is wrong with Bayes' Nets? *The Monist* 84 (2): 242-264, 2001.

[5] Nancy Cartwright, N. Against Modularity, the Causal Markov Condition, and Any Link Between the Two: Comments on Hausman and Woodward. *British Journal for the Philosophy of Science* 53(3): 411-453, 2002.

[6] Nancy Cartwright. Causation: One word; many things. Causality: metaphysics and methods. Technical Report, Centre for the Philosophy of Natural and Social Science, LSE CTR 07-03, 2003.

[7] Nancy Cartwright, N. From Causation To Explanation and Back. In Brian Leiter, editor, *The Future for Philosophy,* pages 230-245. Oxford University Press, 2004.

[8] Nancy Cartwright. From Metaphysics to Method: Comments on Manipulability and the Causal Markov Condition. *British Journal for the Philosophy of Science* 57(1): 197-218, 2006.

[9] L. K. Dennis, L.E. Beane Freeman, M.J. VanBeek. Sunscreen use and the risk for melanoma: a quantitative review. *Annals of Internal Medicine:*139(12):966-78, 2003.

[10] Clark Glymour and Gregory F. Cooper, editors. *Computation, Causation, and Discovery*. AAAI Press, 1999.

[11] Clark Glymour. Rabbit Hunting, *Synthése* 121: 55-78, 1999.

[12] Qingjuan Guan. A *Gaussian Bayesian Network Based Toolkit for Policy Analysis*. MSc Thesis, Department of Computer Science, University of Saskatchewan, 2003.

[13] D. M. Hausman and J. Woodward. Independence, invariance and the Causal Markov Condition. *British Journal for the Philosophy of Science* 50(4): 521-584, 1999.

[14] Daniel M. Hausman and James Woodward. Modularity and the Causal Markov Condition: A Restatement. *British Journal for the Philosophy of Science*, 55(1): 147-161, 2004.

[15] H. E. Kyburg. Causality. *Artificial Intelligence* 169(2): 174-179, 2005.

[16] E. Neufeld, D. Callele, D. Mould, S. Kristtorn, and R. Rabu, A contribution to the theory and practice of cognitive prostheses. In A.Butz, A. Krueger, P. Olivier, editors, *Proceedings of Smart Graphics 2003*, pages 241-250, 2003.

[17] E. Neufeld, S. Kristtorn, Q. Guan, M. Sanscartier, and C. Ware. Exploring Causal Influences. In R.F. Erbacher, J.C. Roberts, editors, *Proceedings of Visualization and Data Analysis 2005*, pages 52-62, 2005.

[18] Eric Neufeld and Sonje Kristtorn. On the Role of the Markov Condition in Causal Reasoning. In *Canadian Conference on AI 2005*, pages 257-267, 2005.

[19] Eric Neufeld and Sonje Kristtorn. Picturing Causality - The Serendipitous Semiotics of Causal Graphs. In *Smart Graphics 2005*, pages 252-262, 2005.

[20] Eric Neufeld and Sonje Kristtorn Whether Non-Correlation Implies Non-Causation. In *FLAIRS Conference 2005*, pages 772-777, 2005.

[21] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, 1988.

[22] J. Pearl and T. Verma. A theory of inferred causation. In J.A. Allen, R. Fikes, and E. Sandewall, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, pages 441-452. Morgan Kaufmann, 1991.

[23] Pearl, Judea. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.

[24] Glenn Shafer. *The Art of Causal Conjecture.* The MIT Press, 1996.

[25] Richard Scheines. An Introduction to Causal inference. In Steven Turner and Vaughan McKim, editors, *Statistical Methods for Causal Knowledge in the Social Sciences*, pages 185-200. Notre Dame, 1997

[26] Bill Shipley. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference*. Cambridge University Press, 2002.

[27] P. Spirtes, C. Glymour and R. Scheines. An Algorithm for Fast Recovery of Sparse Causal Graphs. *Social Science Computer Review* 9: 62-72, 1991.

[28] Peter Spirtes, Clark Glymour and Richard Scheines. *Causation, Prediction, and Search. 2nd Edition.* The MIT Press, 2001.

[29] T. Verma and J. Pearl. An algorithm for deciding if a set of observed independencies has a causal explanation. In D. Dubois, M. P. Wellman, B. D'Ambrosio, and P. Smets, editors, *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, pages 323-330. Morgan Kaufmann, 1992.

[30] Jon Williamson. *Bayesian Nets and Causality: Philosophical and Computational Foundations*, Oxford University Press, 2005.